

# Mesh-to-raster based non-rigid registration of multi-modal images

Rosalia Tatano<sup>a,\*</sup>, Benjamin Berkels<sup>a</sup>, Thomas M. Deserno<sup>b</sup>

<sup>a</sup>Aachen Institute for Advanced Study in Computational Engineering Science (AICES), RWTH Aachen University, Schinkelstraße 2, 52062 Aachen, Germany

<sup>b</sup>Peter L. Reichertz Institute for Medical Informatics, University of Braunschweig and Hannover Medical School, Mühlenpfordtstr. 23, 38106 Braunschweig, Germany

**Abstract.** Region of interest (ROI) alignment in medical images plays a crucial role in diagnostics, procedure planning, treatment, and follow-up. Frequently, a model is represented as triangulated mesh while the patient data is provided from CAT scanners as pixel or voxel data. Previously, we presented a 2D method for curve-to-pixel registration. This paper contributes (i) a general mesh-to-raster (M2R) framework to register ROIs in multi-modal images; (ii) a 3D surface-to-voxel application, and (iii) a comprehensive quantitative evaluation in 2D using ground truth provided by the simultaneous truth and performance level estimation (STAPLE) method. The registration is formulated as a minimization problem where the objective consists of a data term, which involves the signed distance function of the ROI from the reference image, and a higher order elastic regularizer for the deformation. The evaluation is based on quantitative light-induced fluoroscopy (QLF) and digital photography (DP) of decalcified teeth. STAPLE is computed on 150 image pairs from 32 subjects, each showing one corresponding tooth in both modalities. The ROI in each image is manually marked by three experts (900 curves in total). In the QLF-DP setting, our approach significantly outperforms the Insight Segmentation and Registration Toolkit (ITK) mutual information-based registration algorithm.

**Keywords:** Curve-to-pixel registration, Surface-to-voxel registration, ROI alignment, Evaluation, STAPLE, Virtual Physiological Human.

\*Rosalia Tatano, [tatano@aices.rwth-aachen.de](mailto:tatano@aices.rwth-aachen.de)

## 1 Introduction

Two-dimensional (2D) as well as three-dimensional (3D) images play a crucial role in diagnosis, treatment planning, treatment, and the assessment of progression and/or regression of a condition or disease in a patient. In this context, comparisons of subject data acquired using different imaging modalities or of subject and model data are often necessary. Hence, multi-modal image registration methods aim at aligning images (2D) or volumes (3D) acquired with different devices, thus integrating the information provided by this data. The goal is to find the optimal transform that best aligns structures in two input 2D or 3D images.<sup>1</sup>

Mutual information (MI) has been widely used in multi-modal image registration.<sup>2,3</sup> The idea is to maximize iteratively the MI between the two images globally with respect to the transformation. This is equivalent to minimizing the joint entropy of the two data sets, which occurs when the two images or the two volumes are correctly registered.

However, in several cases, there is a particular region of interest (ROI) predefined in the medical recording, such as a tumorous region, a lesion, or some other pathologies. Therefore, accurate ROI alignment is of primary importance. Moreover, in some applications, the ROI might change, e.g., due to pathology or tumor growth. Using MI to register these ROIs might lead to inaccuracies, since MI is computed on the entire image disregarding the corresponding and overlapping parts of the images and hence, is sensitive to the ROI's size and content.<sup>3</sup>

Therefore, ROI-based registration has been addressed in several works already. Wilkie et al.,<sup>4</sup> proposed a modification of MI registration that takes into account information from the ROIs. Yavariabdi et al. presented a registration method for magnetic resonance imaging (MRI) and transvaginal ultrasound (TUS) that matches manually marked contours of the ROIs in the two modalities through a one-step deformable iterative closest point (ICP) method.<sup>5</sup> Gu et al.<sup>6</sup> have proposed a contour-guided deformable image registration algorithm for adaptive radiotherapy that deform images using images intensities and manually marked contours of the ROIs. Finite element model (FEM) -based deformable registration has been employed in Penjweini et al.<sup>7</sup> to match the contours of the ROIs in a series of computed tomography (CT) scans of the lungs acquired pre-operatively with intra-operative images acquired using an infrared camera-based navigation system during the surgery stage of pleural photodynamic therapy.

Whenever considering a ROI, automatic segmentation of such is important. Obviously, the accuracy of the registration depends on the accuracy of segmentation. ROI segmentation is frequently based on the pixel or contour levels, and several approaches have been proposed in the literature. By way of example, the following approaches are taken from dentistry, where digital photography (DP) and quantitative light-induced fluoroscopy (QLF) are common 2D imaging modalities:

- *manual draw*: Hope et al.<sup>8</sup> manually draw the contour around the boundary of the tooth;
- *gray scale*: A threshold technique based on the intensities of the tooth region and of the background is used in Yan et al.<sup>9</sup> to determine the tooth region in fluorescence images;
- *color*: A color based segmentation technique combined with morphological operations is used in Datta et al.<sup>10</sup> to segment a tooth from gum, lips and neighbors teeth in RGB images;
- *statistical model*: In Mansoor et al.,<sup>11</sup> segmentation of the tooth was achieved using a Gauss-Markov random field statistical model;
- *active contour*: In Shah et al.,<sup>12</sup> the contour of teeth from dental images is extracted using an active contour algorithm that depends on the intensity of the tooth region;

Another issue in registration is the multi-modality. Usually, multi-modal means that data from the same subject is taken with different imaging modalities, such as, for example, computed tomography (CT) and magnetic resonance imaging (MRI). Furthermore, atlas registration is required frequently in modern medicine to transfer knowledge coded in a general model (i.e., the atlas) to a specific subject (e.g., the patient in a diagnostic or therapeutic process). Such comparison of virtual physiological human (VPH) models<sup>13</sup> with subject-specific scan data bears another challenge for registration: the VPH models are usually in triangulated mesh-based coding, while patient measurements are obtained with computed axial tomography (CAT) scanners and stored as pixel or voxel data.<sup>14</sup> This yields curve-to-pixel and surface-to-voxel registration problems in 2D and 3D, respectively, disregarding whether registration is considered as global or local (i.e., ROI-based) problem.

In our previous work, a curve-to-pixel-based registration method has been presented<sup>15,16</sup> and used to align the ROIs of 2D images acquired with QLF and DP. The ROIs are segmented using a color space transform into grayscale, which were adapted to the imaging modality, and thresholded. Registration is based on aligning the ROI's contours, i.e., the tooth region. Our curve-to-pixel method allows superimposing DP with QLF and thus, a direct comparison of the detected demineralized areas, an undesirable side effect of orthodontic treatment with fixed appliances.<sup>17</sup>

The aim of the present work is to provide a general registration methodology that is applicable in 2D and 3D; copes with different imaging modalities as well as types of data representations (mesh and raster-scan), and handles global and local (ROI-based) problems. Furthermore, we aim to comprehensively and reliably evaluate the general registration approach using sufficiently large and reliable data sets.

## 2 Materials and methods

In this section, we present a general mesh-to-raster (M2R) registration framework, its application, and the methodology used for quantitative evaluation.

### 2.1 M2R registration method

The M2R method is constructed in the continuous setting as minimization problem and then discretized. After the discretization, we propose a numerical minimization strategy. Moreover, a parametric registration is performed to provide a first alignment of the two input images.

#### 2.1.1 Continuous approach

Let us assume that we are given two data sets, named  $f$  and  $g$ , of the same anatomical structure acquired with different modalities and that the image  $f$  is given as mapping  $f : \Omega \rightarrow \mathbb{R}^n$ ,  $n \in \{2, 3\}$ , on the unit cube  $\Omega = [0, 1]^n$ . Let  $\mathcal{C}$  indicate a hypersurface, i.e. a curve for  $n = 2$  and a surface for  $n = 3$ , representing the contour of the region of interest (ROI), which needs to be extracted from  $g$ , in case  $g$  is also an image or may be identical to  $g$ , if  $g$  was directly acquired as hypersurface, for instance, by a laser scan.

We denote the region of interest in the image  $f$  by  $S_f \subset \Omega$  and now want to find a non-rigid deformation  $\phi : \Omega \rightarrow \mathbb{R}^n$  that matches  $\mathcal{C}$  to the boundary of the set  $S_f \subset \Omega$ . To this end, let  $d$  be the signed distance function of  $S_f$ , i.e.  $d(c, S_f) = \pm \text{dist}(c, \partial S_f)$ . Thus,  $d$  is the Euclidean distance of the point  $c$  to the boundary of  $S_f$ , where the sign is positive if  $c$  is outside of  $S_f$  and negative otherwise.<sup>18</sup> Then, the desired alignment of  $S_f$  and  $\mathcal{C}$  is attained by minimizing the energy

$$E[\phi] = E_{\text{match}}[\phi] + E_{\text{reg}}[\phi] = \frac{1}{2} \int_{\mathcal{C}} w_c (d(\phi(c), S_f))^2 d\mathcal{H}^{n-1}(c) + \frac{\lambda^2}{2} \int_{\Omega} \|\Delta \phi(x)\|^2 dx, \quad (1)$$

where  $\Delta \phi = (\Delta \phi_1, \dots, \Delta \phi_n)$  is the vector of the Laplacian of the components of  $\phi$  and  $d\mathcal{H}^{n-1}$  denotes the  $n - 1$  dimensional Hausdorff measure. Thus, the first term is a hypersurface integral. In particular, it is a curve integral in case  $n = 2$  and a surface integral in case  $n = 3$ .

The energy measures the distance of the deformed hypersurface to the boundary of the ROI in the image  $f$  and the smoothness of the deformation. The parameter  $\lambda > 0$  controls the smoothness of  $\phi$ . Here,  $w_c > 0$  are application dependent weights defined to control the influence of the hypersurface points. Since the data term is using the integral over the hypersurface, it involves only the deformation of  $\mathcal{C}$ . However, the use of a higher order regularizer extends the non-rigid deformation to the whole domain  $\Omega$ .

In contrast to our previous implementation,<sup>15,16</sup> here the data term is formulated as the integral over the hypersurface. Previously, the data term was defined using the sum over the hypersurface points of the weighted signed distance function calculated at the deformed hypersurface points. However, this approach might lead to problems if the hypersurface points are not approximatively

equidistant. Using the integral instead, the distance of a hypersurface point with respect to its neighbors is taken into account.

### 2.1.2 Discretization

The deformation  $\phi$  is expressed as displacement  $u : \Omega \rightarrow \mathbb{R}^n$  via  $\phi(x) = x + u(x)$ , for  $x \in \Omega$ , noting that  $\Delta\phi = \Delta u$ . For the spatial discretization of  $u$ , we use multilinear Finite Elements on a uniform rectangular grid on the image domain  $\Omega$ .<sup>19</sup> Let  $\{\psi_j\}_{j \in J}$  denote the FE basis functions, with nodal index set  $J$ . Let  $M$  and  $L$  denote the lumped mass matrix and the stiffness matrix respectively, i.e.  $M_{i,j} = \int_{\Omega} \mathcal{I}(\psi_i \psi_j) dx$  and  $L_{i,j} = \int_{\Omega} \nabla \psi_i \nabla \psi_j dx$ , where  $\mathcal{I}$  is the bilinear Lagrangian interpolation. Although the chosen regularizer involves second derivatives, it can be approximated as  $E_{\text{reg}}[u] = \frac{\lambda^2}{2} \sum_{i=1}^n \|M^{\frac{1}{2}} L U_i\|^2$ ,<sup>20</sup> using the lumped mass and stiffness matrices. Here,  $U_i$  denotes the vector of nodal values that uniquely represents the scalar FE function  $u_i$ .

### 2.1.3 Minimization

In contrast to our previous implementation,<sup>16</sup> the data term now is evaluated using simplicial finite elements on the hypersurface using a quadrature rule for the numerical evaluation of the integral. The minimization of  $E$  is formulated as nonlinear least squares problem,  $E[u] = \frac{1}{2} \|F[u]\|^2$ , where

$$F[u] = \left[ \left\{ \sqrt{w_c} \sqrt{m_k^q} d(x_k^q + u(x_k^q), S_f) \right\}_{\substack{k=1, \dots, N \\ q=1, \dots, Q}}, \lambda M^{-\frac{1}{2}} L U_1, \dots, \lambda M^{-\frac{1}{2}} L U_n \right]^T. \quad (2)$$

Here,  $m_k^q$  are the weights corresponding to the  $q$ -th quadrature point  $x_k^q$  in the  $k$ -th simplex describing the hypersurface.  $N$  and  $Q$  denote the total number of simplices and of quadrature points in a simplex, respectively. The minimization is efficiently solved using the Gauss-Newton method.<sup>21</sup> To avoid the minimization from getting stuck in local minima, the non-rigid registration problem is solved for decreasing values of the parameter  $\lambda$ . For the sake of simplicity, we use this strategy to avoid a multi level approach, which would also require to create a multi level representation of the unstructured simplicial grid of the hypersurface.

## 2.2 Parametric registration algorithm

To provide a reasonable initial guess for the Gauss-Newton algorithm above, a regularized parametric registration<sup>22</sup> is performed. The aim is to find an affine deformation  $\varphi(c) = Ac + t$ , where  $A$  is a  $n \times n$  matrix and  $t$  a translation vector, that minimizes the energy  $E[\varphi] = E_{\text{match}}[\varphi] + E_{\text{reg}}^{\text{par}}[\varphi]$ , with

$$\begin{aligned} E_{\text{match}}[\varphi] &= \frac{1}{2} \int_{\mathcal{C}} w_c (d(Ac + t, S_f))^2 d\mathcal{H}^{n-1}(c), \\ E_{\text{reg}}^{\text{par}}[\varphi] &:= \frac{\alpha^2}{2} \int_{\Omega} \|J(\varphi(x)) - x\|_F^2 dx + \frac{\mu^2}{2} \int_{\Omega} \sum_{i=1}^n \left\| I - \frac{1}{\partial_{x_i}(\varphi_i(x) - x_i)} D_{J(\varphi(x))} \right\|_F^2 dx \\ &= \frac{\alpha^2}{2} \sum_{i=1}^n \int_{\Omega} \|\partial_{x_i}((A - I)x)\|^2 dx + \frac{\mu^2}{2} \sum_{i=1}^n \left\| I - \frac{1}{a_{ii}} D_A \right\|_F^2, \end{aligned} \quad (3)$$

where  $J(\varphi(x))$  indicates the Jacobian of  $\varphi$  with respect to  $x$  and  $\varphi_i(x)$  represents the  $i$ -th component of the vector  $\varphi(x)$ . Here, the data term is the same as the non-rigid model. In contrast to



our previous formulation,<sup>16</sup> here as prior for the deformation the sum of two terms is used, with positive scalars  $\alpha$  and  $\mu$  that weight the contribution of these terms to the value of the energy. The first term in the regularizer is the Dirichlet energy of the displacement. Noting that  $J(\varphi(x)) = A$ , the second term is the squared Frobenius norm of the matrix  $I - \frac{1}{a_{ii}}D_A$ , where  $I$  is the identity matrix and  $D_A$  is a diagonal matrix whose entries are the diagonal entries of the matrix  $A$ . Using this term, the difference of all the possible ratios of the diagonal entries of  $A$  from 1 is penalized, thus isotropic scalings are preferred to anisotropic ones.

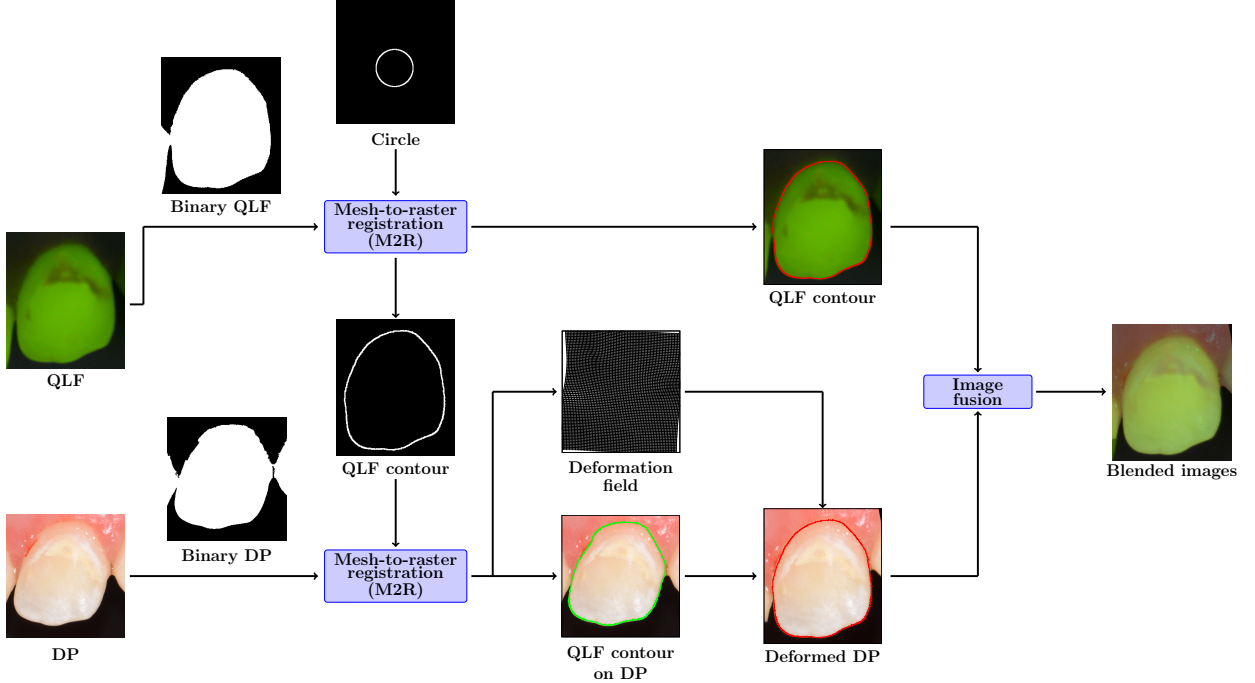
Similarly to the non-rigid case, the parametric registration is formulated as a least squares problem by defining the vector  $F[\varphi]$  and solved using Gauss-Newton.

### 2.3 Applications

In this section, we demonstrate how to apply the method for solving registration problems in 2D and 3D.

#### 2.3.1 2D Example

The selected 2D application aims at registering the tooth as ROI in DP and QLF for demineralization assessment. In contrast to the state of the art in ROI segmentation, we apply the proposed curve-to-image registration method for both the extraction of the tooth contour from the QLF (ROI segmentation) and for its alignment to the tooth region shown in the DP (Fig. 1).



**Fig 1** Workflow of the curve-to-pixel registration method: a DP/QLF pair is given as input; first the algorithm is used to extract the contour of the tooth from the QLF image; then the segmented contour and the DP are used as input for the algorithm, obtaining the segmented DP and a deformation field that is used to align the DP to the QLF. Finally, an image that blends the QLF and the deformed DP can be created.

- *Segmentation*: In the QLF segmentation step, the tooth contour is represented as a small circle in the center of the image, since a tooth is topologically equivalent to a circle. The size of the circle was chosen to ensure that its entire boundary is within the tooth. Then, the registration algorithm aligns the circle to the boundary of the tooth region shown in the QLF, thus, yielding the shape of the tooth.<sup>16</sup> For the alignment of DP and QLF, the proposed algorithm is applied to the extracted QLF tooth contour and the tooth region shown in the DP.
- *Registration*: We assume that the curve is discretized as set of line segments between points  $\{c_i\}_{i=1}^N$ , where  $N$  is the number of points. The data term is discretized using the midpoint rule over each line segment  $[c_i, c_{i+1}]$  as

$$E_{\text{match}}[\phi] = \frac{1}{2} \int_{\mathcal{C}} w_c (d(\phi(c), S_f))^2 dc \approx \sum_{i=1}^N w_{c_i} l_{c_i} \left( d\left(\phi\left(c_{i+\frac{1}{2}}\right), S_f\right) \right)^2 \quad (4)$$

where  $c_{i+\frac{1}{2}}$  indicates the midpoint of the line segment with endpoints  $c_i$  and  $c_{i+1}$  and  $l_{c_i}$  its the length, i.e.  $l_{c_i} = \|c_{i+1} - c_i\|$ . The weights  $w_{c_i}$  of the data term in Eq. (1) are used to get a proper alignment of the curve points on the vertical boundary of the tooth. In fact, these points may have no counterpart in the boundary of  $S_f$ , since often the thresholded images do not exhibit a clear separation between a tooth and its neighbors. Thus, the weights  $w_{c_i}$  are defined using the vector  $\vec{v}_{c_i} = \frac{1}{2}(c_{i+1} - c_i)$  that characterizes the orientation of the curve. The bigger the absolute value of the  $x$ -component of  $\frac{\vec{v}_{c_i}}{|\vec{v}_{c_i}|}$ , the less vertical is  $C$  at  $c_i$ . Hence,  $w_{c_i}$  is set to this value.

Thus, the vector  $F$  (2) that encodes the non-linear least squares problem in the 2D setting is

$$F[u] = \left[ \left\{ \sqrt{w_{c_i} l_{c_i}} d\left(c_{i+\frac{1}{2}} + u\left(c_{i+\frac{1}{2}}\right), S_f\right) \right\}_{i=1, \dots, N}, \lambda M^{-\frac{1}{2}} L U_1, \lambda M^{-\frac{1}{2}} L U_2 \right]^T. \quad (5)$$

The vector  $F$  for the parametric registration step is defined similarly (see Appendix A).

At each iteration of the Gauss-Newton algorithm, the resulting linear system is solved using a sparse Cholesky factorization of the matrix  $J_F^T J_F$ , where  $J_F$  denotes the Jacobian of  $F$ . To this end, we use CHOLMOD from the SuiteSparse<sup>23</sup> library.

### 2.3.2 3D Example

The 3D use-case is taken from the Regional Anaesthesia Simulator and Assistance (RASimAs) project,<sup>24</sup> where subject-specific MRI need to be registered to a VPH model composed of mesh-based surfaces for skin, fascia, muscle, bone, vessels, and nerves.

Hence, we assume that  $f$  is obtained from a volumetric scan, the surface  $\mathcal{C}$  represents the ROI contour of  $g$ , and that  $g$  is given as a triangle mesh. Then, the data term is discretized using linear FE on the triangle mesh as

$$E_{\text{match}}[\phi] = \frac{1}{2} \int_{\mathcal{C}} w_c (d(\phi(c), S_f))^2 dc \approx \sum_{T \in \mathcal{T}} m_T^q (d(\phi(x_T^q), S_f))^2, \quad (6)$$

where the weights  $w_c$  are chosen to be equal to 1,  $T$  is a triangle in the set of triangles  $\mathcal{T}$  defining the triangle mesh of the ROI,  $x_T^q$  is the barycenter of the triangle  $T$  and  $m_T^q$  is the area of  $T$ . Thus, the vector  $F$  (2) that encodes the non-linear least squares problem in the 3D setting is

$$F[u] = \left[ \left\{ \sqrt{m_T^q} d(x_T^q + u(x_T^q), S_f) \right\}_{T \in \mathcal{T}}, \lambda M^{-\frac{1}{2}} LU_1, \lambda M^{-\frac{1}{2}} LU_2, \lambda M^{-\frac{1}{2}} LU_3 \right]^T. \quad (7)$$

The definition of the vector  $F$  for the parametric registration is similar to the non-rigid case. The details are presented in Appendix B.

As in the 2D case, Gauss-Newton requires the solution of a linear system in each iteration of the algorithm. Unlike in the 2D case, it is not feasible though to assemble the system matrix  $J_F^T J_F$  due to memory requirements.  $J_F$  does not only have more rows and columns, but also considerably more non-zero entries in each row. Thus, it is crucial not to assemble this product matrix. Instead, we solve the linear system using the LSMR<sup>25</sup> algorithm, where it is sufficient to compute and store the matrix  $J_F$  and its transposed.

Figure 2 illustrates the effect of the minimization strategy proposed in Section 2.1.3. First the parametric registration is performed with the empirically determined parameters  $\alpha = 1$  and  $\mu = 1$ . Then, the non-rigid registration is performed iteratively for decreasing values of the parameter  $\lambda$ . In this case, the chosen values of the parameters were  $\lambda = 10^{-i}$ ,  $i = \{0, 1, 2, 3, 4\}$ . This allows to get more accurate registration results as  $\lambda$  decreases. For the different settings of  $\lambda$ , Figure 2 visualizes the distances  $d(\phi(c), S_f)$ , for  $c \in \mathcal{C}$ , on the registered template mesh,  $\phi(\mathcal{C}) := \{\phi(c) : c \in \mathcal{C}\}$ , for human hips.

## 2.4 Evaluation

Reliable valuation of non-rigid registration is difficult, since large databases with reliable ground truth (GT) annotations are not available. Nonetheless, we aim at quantitatively evaluating our approach.

### 2.4.1 Aim

Our evaluation aims at determining the absolute error (accuracy) of the method. To define accuracy, we need to rely on a ROI-based registration problem. This requires automatic segmentation, and the accuracy of the segmentation deeply impacts the overall performance. In addition, we aim at comparing the results to a state-of-the-art method for multi-modal registration, which is considered to be based on mutual information (MI).<sup>26</sup>

## 2.5 Metrics

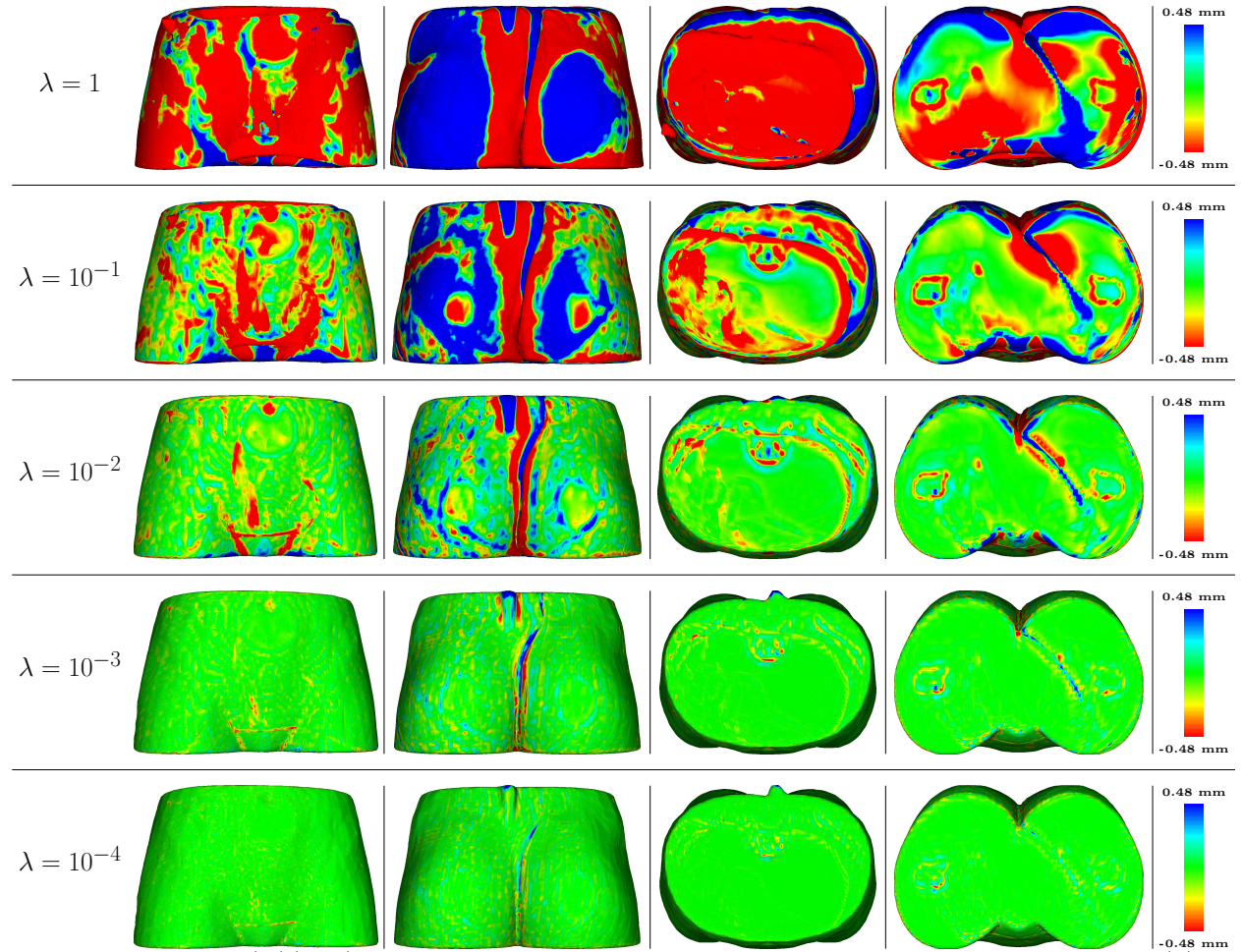
The alignment accuracy is assessed by the Dice Coefficient<sup>27</sup> (DC) and the symmetric Hausdorff distance<sup>28</sup> (HD), which quantify the agreement of two segmentation and the accuracy of the contour alignment, respectively. The larger the DC and the smaller the HD the better the two ROIs correspond.

### 2.5.1 Database

Quantitative measures on ROI alignment are image specific. In order to obtain statistically significant results, a sufficiently large number of images shall be processed. Hence, large databases of images annotated with reliable GT are needed. Therefore, we selected the 2D application of QLF / DP registration, where the ROI is defined as tooth contour. In total, 150 pairs of QLF and DP of upper and lower incisors and canines have been acquired from 32 subjects. All subjects were exhibiting white spot lesions after orthodontic treatment with a fixed appliance.

### 2.5.2 Ground truth

Manually references are unreliable, since they cannot be reproduced exactly, even with the same rater.<sup>29</sup> However, based on several manual markings, a gold standard can be estimated using the simultaneous truth and performance level estimation (STAPLE) algorithm.<sup>30</sup> The core idea of STAPLE is to iteratively (i) compute the observer-weighted mean of a binary segmentation and (ii) adjust the weights of the observers with respect to the similarity to that mean segmentation. In



**Fig 2** The distances  $d(\phi(c), S_f)$ , for  $c \in \mathcal{C}$ , are displayed as color coding on the registered template mesh  $\phi(\mathcal{C})$  for every value of the parameter  $\lambda$  used for the non-rigid registration. From left to right: front, back, top and bottom sides of  $\phi(\mathcal{C})$ , color bar.

other words, if a observer has large discrepancies to the estimated GT, the corresponding weights are lowered in the next iteration.

To generate the ground truth with STAPLE, three trained raters (R1, R2, R3) manually marked the tooth contours on both of the image modalities. All the raters were presented the 300 images in random order.

### 2.5.3 Assessment of segmentation

The accuracy of the automatic ROI segmentation is assessed by calculating the DCs and HDs of the automatic segmentation and the ground truth estimated using the STAPLE algorithm. Including the automatic segmentation, in total four observers are available. The performance of the raters is compared for both modalities, QLF and DP. The contour extraction from the QLF images is done with the parameters  $\alpha = 10^{-3}$ ,  $\mu = 10^{-3}$  and  $\lambda = 10^{-i}$ ,  $i = \{4, 5, 6, 7\}$  for the parametric registration step and the non-rigid step, respectively, while for the DP,  $\alpha = 10^{-3}$ ,  $\mu = 10^{-1}$  and  $\lambda = 10^{-i}$ ,  $i = \{4, 5, 6\}$  were used.

A repeated measures analysis of variance (ANOVA) is applied to assess statistical significances between the automatic segmentation and the human raters, and between the performance of the algorithm in QLF and DP modalities. The significance level is  $\alpha = 0.05$ .

### 2.5.4 Assessment of registration

The accuracy of the DP/QLF ROI alignment is measured again by DCs and HDs of the photo ground truth deformed using the deformation fields obtained from the proposed registration method and the QLF ground truth. For the DP/QLF alignment,  $\alpha = 10^{-3}$ ,  $\mu = 10^{-1}$  and  $\lambda = 10^{-i}$ ,  $i = \{4, 5, 6\}$  were used.

To compare our approach with the state-of-the-art, the same analysis is done using the deformation fields obtained from the Insight Segmentation and Registration Toolkit (ITK)<sup>31</sup> MI registration. The registration is performed on the same grayscale version of the DP/QLF pairs, which are used for the proposed automatic segmentation, using ITK Mattes MI and a third-order B-spline representation for the deformation. The MI is optimized using the Limited-memory Broyden Fletcher Goldfarb Shannon<sup>32</sup> (L-BFGS-B) algorithm.

Based on DC and HD, a one-way repeated measures ANOVA is used to compare the proposed method with the mutual information-based ITK implementation that is considered as gold standard.

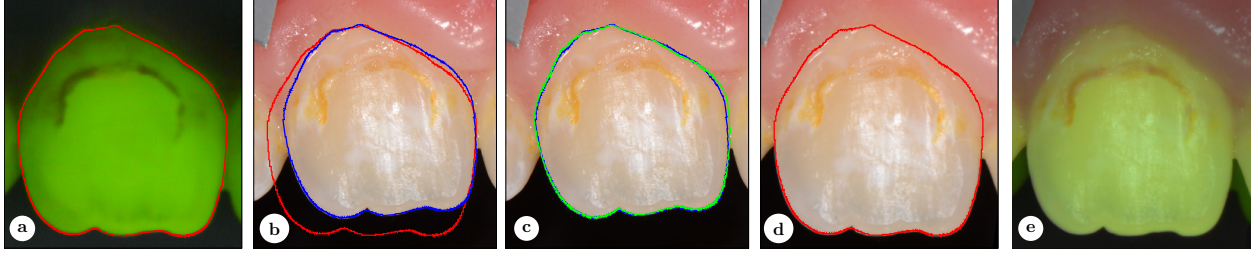
## 3 Results

In this section, we present the results of qualitative registration in 2D and 3D, as well as the accuracy determined in the 2D evaluation for segmentation and registration.

### 3.1 2D application

Figure 3 shows qualitative results obtained in the exemplary 2D application. After non-rigid deformation, the DP matches the QLF, as the QLF-based contour (red) matches the tooth in the DP, as depicted in Panel (d).





**Fig 3** Results in 2D: The QLF image with the the superimposed curve (red) representing the automatic segmentation (a); the DP with superimposed curves (b); QLF-based tooth contour before (red) and after parametric registration (blue); and the curves after parametric and non-rigid registration in blue and green, respectively (c). Panel (d) shows the DP after non-rigid deformation and the corresponding QLF curve (red). Finally, an image created by blending the QLF and the registered DP is displayed in (e).

**Table 1** Mean and standard deviation (SD) of Dice coefficient (DC) and symmetric Hausdorff distance (HD) of the ground truth and manual markings and of ground truth and the proposed automatic segmentation (M2R) for QLF and DP.

Modality	QLF				DP			
Metrics	DC		HD		DC		HD	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Method								
Rater R1	0.990	0.0098	0.015	0.0117	0.984	0.0076	0.021	0.0087
Rater R2	0.980	0.0068	0.023	0.0067	0.991	0.0066	0.015	0.0094
Rater R3	<b>0.992</b>	0.0082	0.012	0.0116	<b>0.996</b>	0.0049	0.008	0.0079
M2R	0.978	0.0101	0.032	0.0185	0.981	0.0083	0.030	0.0129

### 3.2 3D application

The M2R registration method has been applied in 3D, too. The resolution of the corresponding 3D grid is  $129 \times 129 \times 129$ . Figure 4 depicts the results and suggests an appropriate surface-to-voxel alignment.

### 3.3 Quantitative evaluation of segmentation performance

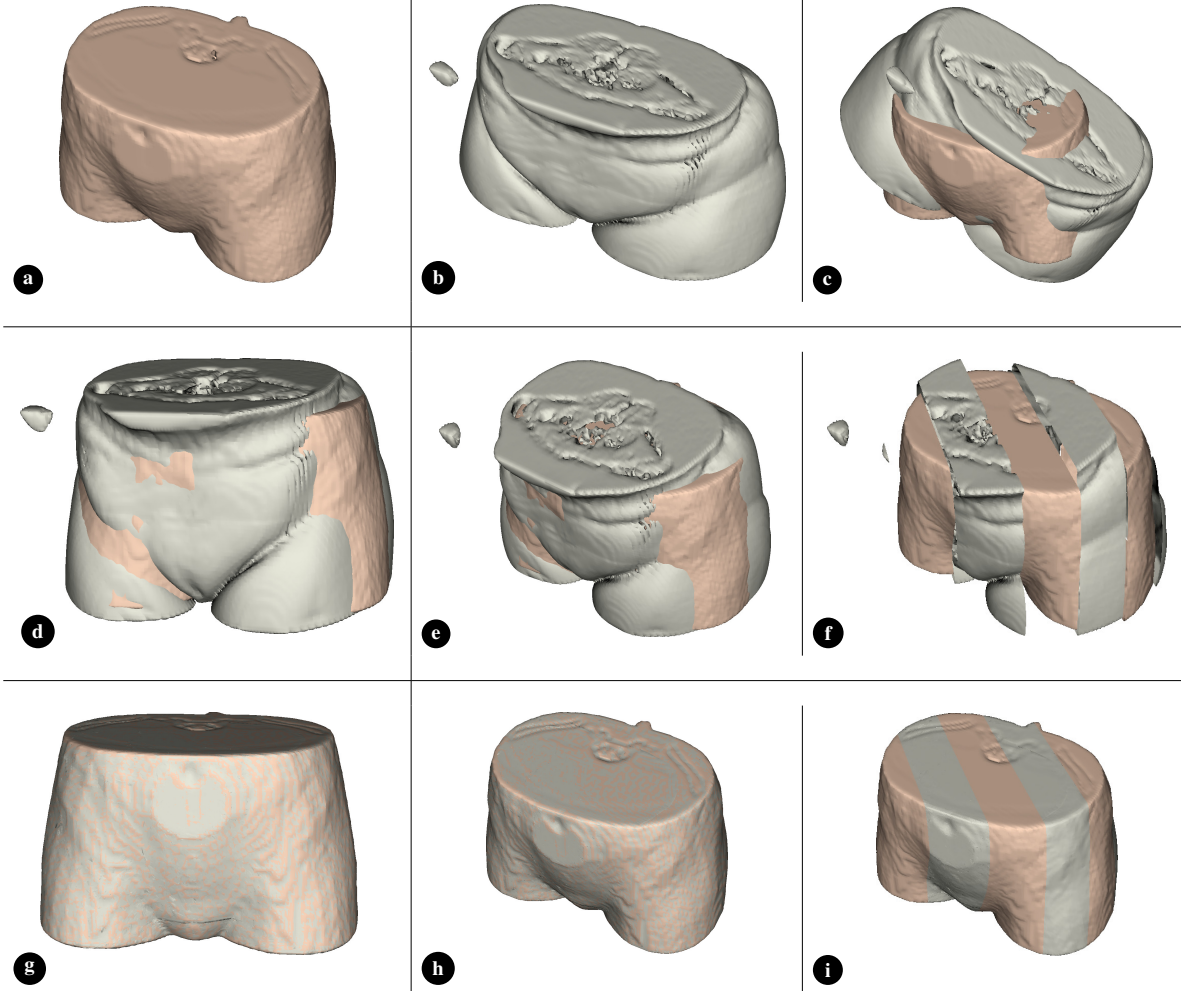
Based on the 150 QLF images, the mean of the DC between the segments obtained from the automatic procedure and the STAPLE-based GT that is calculated from the three human observers is 0.978 (range 0.9420-0.9940). The mean HD is 0.032 (range 0.0102-0.1060).

Based on the 150 DP images, these values turn to 0.981 (range 0.9440-0.9940) and 0.030 (range 0.0102-0.0792) for DC and HD, respectively (Table 1).

For both QLF and DP, a pairwise comparison of the means of the DC as well as of the HD using a one-way repeated measure ANOVA was performed between the automatic segmentation and each rater, determining a statistical significant difference between these means (Table 3).

The one-way repeated measures ANOVA revealed a statistical significant difference between the means of the DC ( $F(1, 149) = 8.68, p = 0.0037$ ) between DP and QLF modalities, but no significant difference between the means of the HD ( $F(1, 149) = 1.86, p = 0.1742$ ).





**Fig 4** Results in 3D. The first row shows the starting setting:  $S_f$  (a), template mesh  $g$  (b), and initial position of the input images (c). Note that (a) and (b) use different manually chosen view angles to simplify the comparison, (c) shows the true initial mismatch of the data sets. Middle and lower row depict the results after parametric registration and non-rigid deformation, respectively. Panels (f) and (j) visualize the according slices of  $S_f$  and of the registered template mesh after parametric and non-rigid registration respectively.

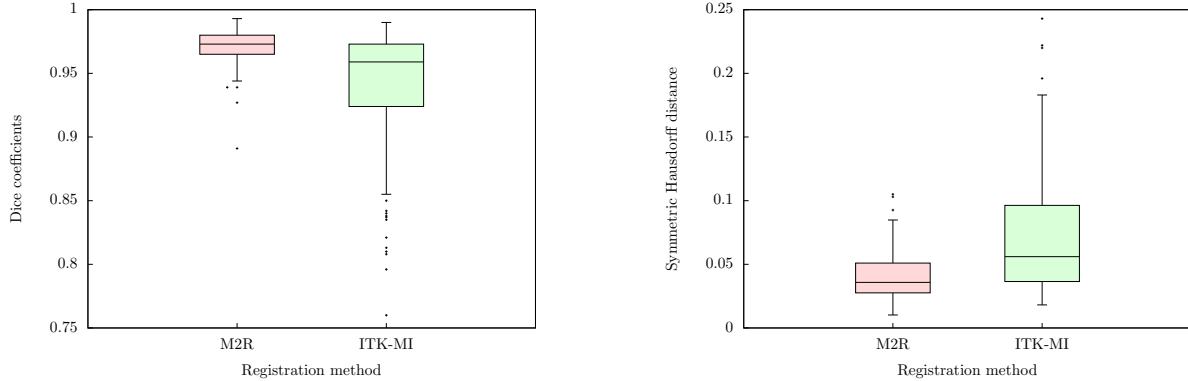
### 3.4 Quantitative evaluation of registration performance

The descriptive statistics (mean and standard deviation) of DCs and HDs measuring the accuracy of the alignment for the ITK-MI state-of-the-art method and our mesh-to-raster (M2R) approach are reported in Table 2. Based on the 150 QLF/DP image pairs, the means of the DC are 0.940 (range 0.7600-0.9900) and 0.971 (range 0.8910-0.9930) for ITK-MI and M2R, respectively. The means of the HD is 0.072 (range 0.0181-0.2630) and 0.041 (range 0.0102-0.1050), respectively.

The one-way repeated measures ANOVA revealed a statistical significant difference between the means of the DC ( $F(1, 149) = 71.96, p < 0.001$ ) as well as the means of the HD ( $F(1, 149) = 61.99, p < 0.001$ ). The statistical significance is emphasized by the box plots in Figure 5, which are visualizing the data from Table 2. Both the DCs and the HDs exhibit more variability in the case of ITK-MI rather than for our method.

**Table 2** Mean and standard deviation (SD) of Dice coefficient (DC) and symmetric Hausdorff distance (HD) of the deformed DP ground truth and the QLF ground truth for the proposed mesh-to-raster (M2R) and the ITK-MI reference methods.

Metrics	DC		HD	
	Mean	SD	Mean	SD
<b>Method</b>				
ITK-MI	0.940	0.0471	0.072	0.0494
M2R	<b>0.971</b>	0.0129	<b>0.041</b>	0.0180



**Fig 5** Boxplots of the Dice coefficients and the symmetric Hausdorff distance of the QLF ground truth and the DP ground truth deformed using the proposed algorithm (M2R) and the ITK mutual information (ITK-MI).

## 4 Discussion

The alignment of ROIs in multi-modal medical images is important for many applications. Here, a mesh-to-raster (M2R)-based method is described, which aligns the contour of the ROI of the reference image to the target image with a deformation field that is determined on the whole image domain. Although this paper provides examples in 2D and 3D only, the method is applicable in general in any dimension.

The quantitative evaluation is based on multi-modal 2D data from dentistry. Our methods was applied to QLF and DP images, including an evaluation of the automatic segmentation that is needed to determine the tooth region, which acts as ROI.

With respect to the evaluation of the ROI segmentation, our M2R approach is not as accurate as a human observer (a significant difference among the means of each human rater and the automatic segmentation was determined by ANOVA). In absolute numbers for QLF, the distance in terms of mean DC between M2R (DC = 0.978) and the best-performing human observer R3 (DC = 0.992) is only about 1.5%, while the distance between the M2R and the worst-performing human observer R2 is negligible (about 0.2%). Similar results hold for the DP. Also, the mean of the HD for the automatic segmentation is slightly bigger than those of the manual markings.

The observed deviation from the GT in a QLF is due to the extended and smooth transition zone between tooth and background, leading to an imprecise classification of the image into tooth and non-tooth regions and thus to a less accurate contour extraction than in cases where the images exhibit clear distinction between tooth and background. Similarly for the DP images, small inaccuracies in tooth and non-tooth regions classification are caused mostly by the low contrast

QLF with marked GT	DP with marked GT	DP deformed using M2R	DP deformed with ITK-MI	Quality measure for M2R	Quality measure for ITK-MI
				HD (0.0397) DC (0.9750)	Best HD (0.0181) DC (0.9820)
				Best DC (0.9930) Best HD (0.0102)	Best DC (0.9900) HD (0.0201)
				DC (0.9760) HD (0.0322)	Worst DC (0.7600) Worst HD (0.2630)
				Worst DC (0.8910) Worst HD (0.1050)	DC (0.8810) HD (0.1190)

**Fig 6** Best and worst DC and HD for the proposed method (M2R) and the ITK-MI. From left to right: QLF with marked GT contour (red), DP with marked GT contour (white), DP with marked GT contour (blue) deformed using the M2R and contour extracted from QLF (red) and DP with marked GT contour (blue) deformed using ITK mutual information (ITK-MI) and contour extracted from QLF (red).

between the colors of tooth and gum or adjacent teeth. However, the segmentation step may be replaced easily by another algorithm.

Contour-based registration methods have been used in several works for ROI extraction and alignment. Chen et al.<sup>33</sup> used an automatic method to extract and align teeth contours to register dental radiographs. A deformable image registration method that used ROI's contour propagation was proposed by Wu et. al<sup>34</sup> in radiotherapy. Here, the proposed method is used for both ROI extraction and matching. The method is correctly aligning the contour of the ROI in the target image to the bounday of the ROI in the reference image. However, if the boundary of ROI in the reference image is not accurately delineated, it might lead to inaccuracies in the registration step, as already pointed out in the case of QLF and DP segmentation.

The best and worst DCs and HDs for both our method and ITK-MI registration are depicted in Fig. 6. These results clearly illustrate the improved accuracy of the proposed method over ITK-MI that was shown quantitatively but in an abstract manner by the DC and HD means. Disregarding that we have used a non-optimal ROI extractor, we have been able to outperform the ITK-MI registration significantly. This is due to the ROI-based vs. global approach, respectively. It emphasizes the need of ROI-based registration, which is in line with Yi et al.<sup>35</sup>

The selection of ITK-MI as reference methods is a limitation of our evaluation. Since MI has

been considered as state-of-the-art about 15 years ago,<sup>3</sup> several approaches have been published. However, multi-modal and/or ROI-based approaches are rarely found already implemented.<sup>1</sup>

Another limitation of our work is that the quantitative evaluation is performed only in 2D. In future, we plan to obtain quantitative assessment in 3D, too. Here data from the Evaluation of Methods for Pulmonary Image Registration 2010 (EMPIRE10) or other challenges might be helpful.<sup>36</sup> In this context, the proposed method could be used to perform lung registration by aligning the lung boundaries. However, to align the major fissures inside the lung, an additional data term would need to be included in the proposed method, which we plan to exploit in a future work.

## 5 Conclusions

A ROI-based registration method for multi-modal images was presented. It uses a curve-to-pixel or surface-to-voxel approach to align the ROIs from the reference and the target image in 2D or 3D, respectively. Qualitative examples in 2D and 3D were presented. The accuracy of the alignment was tested on multi-modal 2D images. Ground truth was established by applying the STAPLE algorithm to manually marked images. In a comprehensive analysis based on 150 pairs of images, the proposed method statistical significantly outperforms mutual information (MI)-based global registration, which is considered as state-of-the-art method.<sup>3</sup>

## Appendix A: details of the parametric registration in 2D

For the parametric registration in 2D, if  $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ ,  $D_A = \begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix}$ , the regularizer can be expressed as

$$E_{\text{reg}}[\varphi] = \frac{\alpha^2}{2} ((a_{11} - 1)^2 + a_{12}^2 + a_{21}^2 + (a_{22} - 1)^2) + \frac{\mu^2}{2} \left( \left(1 - \frac{a_{11}}{a_{22}}\right)^2 + \left(1 - \frac{a_{22}}{a_{11}}\right)^2 \right).$$

Therefore, the vector  $F$  for the parametric registration is

$$F[\varphi] = \begin{bmatrix} \left\{ \sqrt{w_{c_i} l_{c_i}} d \left( \varphi \left( c_{i+\frac{1}{2}} \right), S_f \right) \right\}_{i=1, \dots, N} \\ \alpha(1 - a_{11}) \\ \alpha a_{12} \\ \alpha a_{21} \\ \alpha(1 - a_{22}) \\ \mu \left( 1 - \frac{a_{11}}{a_{22}} \right) \\ \mu \left( 1 - \frac{a_{22}}{a_{11}} \right) \end{bmatrix}.$$

## Appendix B: details of the parametric registration in 3D

For the parametric registration in 3D, if  $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ ,  $D_A = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$ , the regularizer for the parametric registration, (3), can be expressed as

$$E_{\text{reg}}^{\text{par}}[\varphi] = \frac{\alpha^2}{2} ((a_{11} - 1)^2 + a_{12}^2 + a_{13}^2 + a_{21}^2 + (a_{22} - 1)^2 + a_{23}^2 + a_{31}^2 + a_{32}^2 + (a_{33} - 1)^2) \\ + \frac{\mu^2}{2} \left( \left(1 - \frac{a_{11}}{a_{22}}\right)^2 + \left(1 - \frac{a_{11}}{a_{33}}\right)^2 + \left(1 - \frac{a_{22}}{a_{11}}\right)^2 + \left(1 - \frac{a_{22}}{a_{33}}\right)^2 + \left(1 - \frac{a_{33}}{a_{11}}\right)^2 + \left(1 - \frac{a_{33}}{a_{22}}\right)^2 \right).$$

Therefore, the vector  $F$  for the parametric registration is

$$F[\varphi] = \begin{bmatrix} \left\{ \sqrt{m_T^q} d(x_T^q + u(x_T^q), S_f) \right\}_{T \in \mathcal{T}} \\ \alpha(1 - a_{11}) \\ \alpha a_{12} \\ \alpha a_{13} \\ \alpha a_{21} \\ \alpha(1 - a_{22}) \\ \alpha a_{23} \\ \alpha a_{31} \\ \alpha a_{32} \\ \alpha(1 - a_{33}) \\ \mu \left(1 - \frac{a_{11}}{a_{22}}\right) \\ \mu \left(1 - \frac{a_{11}}{a_{33}}\right) \\ \mu \left(1 - \frac{a_{22}}{a_{11}}\right) \\ \mu \left(1 - \frac{a_{22}}{a_{33}}\right) \\ \mu \left(1 - \frac{a_{33}}{a_{11}}\right) \\ \mu \left(1 - \frac{a_{33}}{a_{22}}\right) \end{bmatrix}.$$

## Appendix C: details of the ANOVA for the segmentation evaluation

This section presents the results ( $F$  and  $p$ -value) of the one-way repeated measures ANOVA for the comparison of the means of the Dice coefficient (DC) as well as of the symmetric Hausdorff distance (HD) of each human rater and the automatic segmentation for both the QLF and the DP.

**Table 3** ANOVA results for the comparison of the means of the Dice coefficient (DC) and symmetric Hausdorff distance (HD) of the human raters and the automatic segmentation for QLF and DP.

Modality	QLF				DP			
	DC		HD		DC		HD	
	$F(1, 149)$	$p$	$F(1, 149)$	$p$	$F(1, 149)$	$p$	$F(1, 149)$	$p$
(R1,M2R)	113.16	< 0.001	107.73	< 0.001	20.01	< 0.001	54.85	< 0.001
(R2,M2R)	8.76	0.0036	38.34	< 0.001	155.08	< 0.001	139.09	< 0.001
(R3,M2R)	234.00	< 0.001	134.41	< 0.001	415.61	< 0.001	345.48	< 0.001

## Informed consent

Informed consent was obtained from all individuals participating in the trial. The study was approved by the Ethics Committee of the RWTH Aachen University.

## Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

## Acknowledgments

The authors would like to thank Eva E. Ehrlich and Ekaterina Sirazitdinova, Uniklinik RWTH Aachen, for their assistance in the preparation of the 2D data. The authors at AICES RWTH Aachen were funded in part by the Excellence Initiative of the German Federal and State Governments. The RASimAs project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 610425. The authors declare no conflict of interest and have nothing to disclose.

## References

- 1 A. P. Keszei, B. Berkels, and T. M. Deserno, "Survey of non-rigid registration tools in medicine," *J Digit Imaging*, 1–15 (2016).
- 2 P. Viola and W. M. Wells, "Alignment by maximization of mutual information," *Int J Comput Vision* **24**(2), 137–154 (1997).
- 3 J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Trans Med Imaging* **22**(8), 986–1004 (2003).
- 4 K. P. Wilkie and E. R. Vrscay, "Mutual information-based methods to improve local region-of-interest image registration," in *Proceedings of the Second International Conference on Image Analysis and Recognition, ICIAR'05*, 63–72, Springer-Verlag, (Berlin, Heidelberg) (2005).
- 5 A. Yavariabdi and et al., "Contour-based TVUS-MR image registration for mapping small endometrial implants," in *Abdominal Imaging. Computation and Clinical Applications: 5th International Workshop, Held in Conjunction with MICCAI 2013, Nagoya, Japan, September 22, 2013. Proceedings*, H. Yoshida, S. Warfield, and M. W. Vannier, Eds., 145–154, Springer Berlin Heidelberg, (Berlin, Heidelberg) (2013).
- 6 X. Gu and et al., "A contour-guided deformable image registration algorithm for adaptive radiotherapy," *Phys Med Biol* **58**(6), 1889 (2013).
- 7 R. Penjweini and et al., "Deformable medical image registration of pleural cavity for photodynamic therapy by using finite-element based method," in *Proc SPIE Int Soc Opt Eng. 9701*, 970106 (2016).
- 8 C. Hope and et al., "Assessing the association between oral hygiene and preterm birth by quantitative light-induced fluorescence," *Scientific World J* **2014**(374694), 1–9 (2014).
- 9 J. Yan, Y. Xiang, and X. Jian, "Automated detection and quantification of early caries lesions on images captured by intraoral camera," in *Bioelectronics and Bioinformatics (ISBB), 2011 International Symposium on*, 251–254 (2011).



- 10 S. Datta and N. Chaki, "Detection of dental caries lesion at early stage based on image analysis technique," in *2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS)*, 89–93 (2015).
- 11 A. Mansoor and et al., "A statistical modeling approach to computer-aided quantification of dental biofilm," *IEEE J Biomed Health Inform* **19**(1), 358–366 (2015).
- 12 S. Shah and et al., "Automatic tooth segmentation using active contour without edges," in *2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, 1–6 (2006).
- 13 M. Viceconti and P. Hunter, "The virtual physiological human: Ten years after," *Annu Rev Biomed Eng* **18**(1) (2016).
- 14 J. E. E. de Oliveira and et al., "Surface mesh to voxel data registration for patient-specific anatomical modeling," *Proc. SPIE* **9786**, *Medical Imaging 2016: Image-Guided Procedures, Robotic Interventions, and Modeling*, 978625 (2016).
- 15 B. Berkels and et al., "Non-rigid contour-to-pixel registration of photographic and quantitative light-induced fluorescence imaging of decalcified teeth," *Proc. SPIE* **9784**, *Medical Imaging 2016: Image Processing*, 97840Z–7 (2016).
- 16 B. Berkels and et al., "Curve-to-image based non-rigid registration of digital photos and quantitative light-induced fluorescence images in dentistry," in *Bildverarbeitung für die Medizin 2016, Informatik aktuell*, 80–85, Springer (2016).
- 17 L. Gorelick, A. Geiger, and A. Gwinnett, "Incidence of white spot formation after bonding and banding," *Am J Orthod* **81**(2), 93–98 (1982).
- 18 S. Bauer and et al., "Joint ToF image denoising and registration with a CT surface in radiation therapy," in *Lect Notes Computer Sci*, 98–109 (2011).
- 19 D. Braess, *Finite Elements: Theory, Fast Solvers, and Applications in Solid Mechanics 3rd Edition*, Cambridge University Press (2007).
- 20 B. Berkels and et al., "Co-registration of intra-operative brain surface photographs and pre-operative MR images," *Int J Comput Assist Radiol Surg* **9**(3), 387–400 (2014).
- 21 S. Gratton, A. S. Lawless, and N. K. Nichols, "Approximate Gauss-Newton methods for nonlinear least squares problems," *SIAM J Optim* **18**(1), 106 – 132 (2007).
- 22 N. Chumchob and K. Chen, "A robust affine image registration method," *Int J Num Anal Model* **6**(2), 311–334 (2009).
- 23 Y. Chen and et al., "Algorithm 887: CHOLMOD, supernodal sparse Cholesky factorization and update/downdate," *ACM Trans Math Softw* **35**(3), 22:1–22:14 (2008).
- 24 T. Deserno, J. Oliveira, and O. Grottke, "Regional anaesthesia simulator and assistant (RASi-mAs): Medical image processing supporting anaesthesiologists in training and performance of local blocks," in *Proceedings of the 28th IEEE International Symposium on Computer-Based Medical Systems*, 348–51 (2015).
- 25 D. C.-L. Fong and M. Saunders, "LSMR: an iterative algorithm for sparse least-squares problems," *SIAM J Sci Comput* **33**(5), 2950–2971 (2011).
- 26 F. P. Oliveira and J. M. R. Tavares, "Medical image registration: a review," *Comput Methods Biomech Biomed Engin* **17**(2), 73–93 (2014).
- 27 L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology* **26**(3), 297–302 (1945).

- 28 D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans Pattern Anal Mach Intell* **15**(9), 850–863 (1993).
- 29 T. Lehmann, "From plastic to gold: a unified classification scheme for reference standards in medical image processing," *Proc SPIE* **4684**(3), 1819–27 (2002).
- 30 S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation," *IEEE Trans Med Imaging* **23**(7), 903–921 (2004).
- 31 H. J. Johnson, M. M. McCormick, and L. Ibanez, *The ITK Software Guide: Introduction and Development Guidelines*, Kitware Inc., fourth ed. (2015).
- 32 C. Zhu, R. H. Byrd, and J. Nocedal, "L-BFGS-B: Algorithm 778: L-BFGS-B, FORTRAN routines for large scale bound constrained optimization," *ACM Trans Math Softw* **23**(4), 550–560 (1997).
- 33 H. Chen and A. K. Jain, "Dental biometrics: alignment and matching of dental radiographs," *IEEE Trans Pattern Anal Mach Intell* **27**(8), 1319–1326 (2005).
- 34 Q. Wu and et al., "Deformable image registration of CT images for automatic contour propagation in radiation therapy," *Biomed Mater Eng* **26**, S1037–S1044 (2015).
- 35 W. Yi and et al., "ROI-based image registration for digital subtraction radiography," *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* **101**, 523–529 (2006).
- 36 K. Murphy and et al., "Evaluation of registration methods on thoracic CT: the EMPIRE10 challenge," *IEEE Trans Med Imaging* **30**(11), 1901–1920 (2011).